DOCUMENT RESUME

ED 147 844 CS 203 799

AUTHOR Reich, Carol M.; Reich, Peter M.
TITLE The Vocabulary of Primary School Children.
INSTITUTION Toronto Board of Education (Ontario). Research Dept.
SPONS AGENCY Ontario Dept. of Education, Toronto.; Ontario Inst. for Studies in Education, Toronto.
PUB DATE May 77
NOTE 45p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Basic Reading; *Beginning Reading; *Child Language; Foreign Countries; Language Development; Language Research; *Language Usage; Primary Education; *Reading Research; *Vocabulary; *Word Lists
IDENTIFIERS Ontario (Toronto)

ABSTRACT

This study investigated the agreement between previously published vocabulary lists recommended for use with children and lists derived from the speech of contemporary children in grades one, two, and three in Toronto. In the first part of the study, existing vocabulary lists were compared. In the second part, language samples were collected from four diverse groups of Toronto children and new vocabulary lists were constructed; these were then compared with previous lists and with the words used in several beginning reading series. In the third part of the study, the language samples were analyzed for linguistic complexity, in order to examine developmental differences between the four groups. Among the findings were that the oral vocabulary of children is quite different from vocabulary in material written for children and that the oral vocabulary of children is quite sensitive to the manner in which it is collected. With the methods used in this study, however, economically advantaged and disadvantaged children, as well as those from non-English backgrounds, were equally talkative and had equally rich vocabularies. A new vocabulary list based on the Toronto study is included in the report, and it is suggested that this list is suitable for use with a wide variety of children beginning to read. (RL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED147844

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ISSN 0316-8786

THE VOCABULARY OF PRIMARY SCHOOL CHILDREN

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Toronto Board of Education

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Carol M. Reich*

Peter M. Reich**

#144

May, 1977

CS203 799

This research was supported by a Grant-in-Aid of Educational Research from the Ministry of Education in 1972, by the Research Department of the Toronto Board of Education, where the first author was Research Associate at the time, and by a small scale research grant from the Ontario Institute for Studies in Education in 1975-76.

* Ontario Institute for Studies in Education.

** University of Toronto.

ACKNOWLEDGMENTS

This study owes its existence to Mr. Wilson Brooks, principal of one of the schools in the study, who first suggested the idea and worked with the first author in developing the original proposal.

That school participated in the research, along with three others, and their principals and teachers also deserve our thanks for working with us in developing procedures and then collecting the language samples from the children.

Three graduate students in Special Education at the Ontario Institute for Studies in Education worked on various stages of the data analysis: Mrs. Kathy Margaritis, Mrs. Thiruvambigal Harris, Miss Tswei-Hwa Lee.

This research was supported by three agencies: the Ministry of Education in their Grants-in-Aid of Educational Research, the Research Department of the Toronto Board of Education, and small scale research funds from the Ontario Institute for Studies in Education.

TABLE OF CONTENTS

	Page
Introduction	1
Which List to Select: A Comparison of Vocabulary Lists	4
A Study of the Vocabulary of Toronto Children	10
Analysis of the Toronto Lists and Comparison With Previous Lists	17
A Toronto List	23
A Comparison of the Vocabulary in Beginning Readers With the Toronto and non-Toronto Lists	23
A Comparison of the Toronto Samples on Grammatical Measures	25
Summary and Abstract	33
REFERENCES	36
APPENDIX A	38

THE VOCABULARY OF PRIMARY SCHOOL CHILDREN

Introduction

There have been many attempts to construct vocabulary lists for children in the early grades. The rationale for constructing such lists is that the use of a controlled vocabulary makes learning to read easier.

One argument put forward is that it will be easier for children to learn words with which they are familiar. The most straightforward way of assessing familiarity is to collect samples of children's speech and determine the frequency with which various words appear. Carroll (1972) summarizes research which shows that word frequency relates to familiarity with a word, knowledge of its meaning, and speed of recognition.

Most lists, however, are based on an analysis of the frequency with which words appear in written materials, often even materials written for adults. One obvious reason for using written materials is that it is simpler to analyze existing texts than to collect speech samples and begin the research from there. A likely reason for using adult materials is that there is available large quantities of different types of materials, making possible a comprehensive study of English usage. Large-scale vocabulary studies of adult materials can serve several purposes, while an analysis of materials written for children have a more narrowly educational focus.

Carroll, however, suggests that the frequency with which words occur in written materials may not reflect spoken usage, and, therefore, that such lists may not be a good reflection of what

children actually know. Hillerich (1974) contends that there is even a difference between the language spoken and the language written by children, and an even greater difference between children's spoken language and language written for them by adults. There would, no doubt, be a still greater difference between children's spoken language and adult reading materials.

A different point of view is that vocabulary lists should reflect the words that children are most likely to encounter in printed materials. There is no point in teaching children to read a list of words if there are no materials for them to read (Johnson, 1971). Following this line of reasoning some researchers have constructed lists from readers and textbooks prepared for children.

However, even accepting this viewpoint, it can still be argued that word lists constructed from oral language are useful as a guide to authors. They might also be useful for comparing with the vocabulary in existing reading materials as an aid to material selection.

No one would argue that vocabulary should be the only criterion used by curriculum planners. It is possible that some reading materials which use highly esoteric vocabulary are most appropriate for children because they have high interest value. Children may not talk about ghosts and goblins very much, but they might like to read about them. Educators would also want to consider the literary value of stories when selecting materials.

Finally, it is also important to keep in mind that vocabulary analysis is only the first step in a study of the appropriateness of the language used in any reading materials. It is possible for a reading series to use very common words, but combine them in uncommon

grammatical structures. Such a reading series would also present unnecessary difficulties for the student just learning to read.

This study, then, is based on the premise that a list of words most commonly used by children is a useful tool for the curriculum planner. Congruence between this list and the words used in a reading series for children is one criterion to be used in material selection or creation. It is not the sole or even the most important criterion, but it is one piece of information which should be taken into account.

The study reported here has three parts. In the first part, existing vocabulary lists were compared in order to determine whether or not there were important differences among them. In the second part of the study, language samples were collected from four groups of Toronto children and new vocabulary lists constructed. The resulting lists were then compared with previous lists and with the words used in several beginning reading series. The purpose of this was to determine how well the various series reflected the word usage of a contemporary group of local children.

In a third part to the study, the language samples collected from Toronto children were analyzed for linguistic complexity in order to determine if there were any developmental differences among the four groups. This was a small pilot study meant to suggest what other special needs particular groups of children might have with respect to the language and reading curriculum of the school.

Which List to Select: A Comparison of Vocabulary Lists

A recent study by Harris and Jacobson (1974) suggests that it does not matter which list one selects, since there is very little difference among them. In this study Harris and Jacobson compared 13 vocabulary lists, including their own. They found that their own list embraced almost all of the words occurring on the other lists, and they therefore concluded that there is a high level of agreement among various lists. However, their conclusions can be challenged on two grounds.

First of all, 12 of the lists were based mainly on an analysis of written material for children or adults. In some cases, the analysis of written material was supplemented by a study of the actual speech of children. But these studies of children's speech were either quite old or they were not done in Canada. The 13th list comprised 3000 words that fourth grade children demonstrated that they knew. However, the words selected for testing had been chosen from a list based on written materials, and the study was done in the United States in 1948. Therefore the results may not reflect important vocabulary items in the spontaneous spoken language of contemporary Canadian children.

A second problem with the study is that Harris and Jacobson compared relatively small lists with their own large list of more than 16,000 entries. One of the lists they used contains almost 7000 items; 3 lists are in the 3000 item range, and 6 lists have fewer than 800 entries. They did include two lists that had more than 50,000 entries; but they drew small samples of words from these lists and compared them with larger portions of their own

lists. It is not at all surprising that the words occurring on these small lists almost always occurred on the large comparison list as well.

The Harris and Jacobson list also assigns a grade level to each word, according to its occurrence in children's texts. Harris and Jacobson found that different lists had drawn words from different reading levels. They also found that different lists disagreed on the relative frequencies of specific words. They considered this of minor importance. Simple presence or absence of a word was their primary criterion of agreement. But relative placement of a word on a list is of prime importance when a relatively short list is being constructed for a sight word vocabulary, or when instructional materials for children are being selected. The most that we can conclude from the Harris and Jacobson study is that there is a great deal of overlap among large vocabulary lists. But this is not too surprising since there is a limit to the number of words in the English language. Even two random selections of words would have a great deal of overlap if they were large enough.

One of the more recently constructed lists is the Great Atlantic and Pacific Sight Word List developed by Otto and Chester (1972). This list consists of the 500 most frequent words from a corpus of grade 3 reading materials. Aware of the possible bias resulting from written materials, Otto and Chester (1974) later compared their list with the 1000 word Newman-Bailey list. This latter list consists of the 1000 most frequent words from a composite of eight frequency counts of the oral language of pre-school and kindergarten children. Otto and Chester found that 389 or 78 per cent of the 500 words on the Atlantic and Pacific list appeared on the 1000 word Newman-Bailey (1971) list. They concluded that the two lists were very similar, and

that "the difference in children's spoken language and the language used in school texts may be more fancied than factual."

However, in their original article, Otto and Chester suggest that even a 500 word list is too long. Certainly, then 1000 is too many. It is therefore more appropriate to compare the Atlantic and Pacific list with only the first 500 words of the Newman-Bailey list. When this is done, it is found that only 317 or 62 per cent of the words are held in common.

The most well-know of all vocabulary lists is the one developed by Dolch (1941). The Dolch List actually consists of two lists. The Dolch Basic List contains 220 words which Dolch felt were context-free, and therefore would be encountered by a child in a wide variety of reading materials. This list includes words in all grammatical categories except the noun. Dolch also prepared a second list containing 95 nouns which he thought were widely occurring.

The Dolch list is the only vocabulary list to make a distinction between types of words, and it seems a useful thing to do. However, one can question whether Dolch made this distinction in the most reasonable way. Adjectives, adverbs, and verbs appear on Dolch's Basic List, and it can be argued that such words are restricted to certain contexts although not to the same extent as nouns.

It would seem more reasonable to us to include adjectives, adverbs, and verbs with nouns in the category of "content" words, i.e. words which describe the topic under discussion. All other words -- i.e. pronouns, articles, prepositions, auxiliaries, the

copula and conjunctions -- are "function" words which specify relationships among the components of a sentence.

Another reason for drawing the distinction in this way is the fact that nouns, verbs, adjectives, and adverbs are "slippery" with respect to grammatical categories. The same root word can be both a noun and verb -- e.g., play, ship, ground, cover -- or a noun and an adjective -- e.g., gold, glass, paper. Furthermore, English contains morphemic devices for transforming verbs into nouns - e.g., act and action, run and running -- and adjectives into adverbs -- e.g., beautiful and beautifully, pretty and prettily. Without delineating the context, it is often impossible to specify to which category a content word belongs.

Applying this simple distinction to the Atlantic and Pacific and Newman-Bailey lists, it becomes clear that there is a great deal of similarity in the occurrence of function words, but less similarity in content words. Otto and Chester mention that 95 of the first 100 words on both lists are the same. However, 55 of these are function words, all of which appear on both lists. Of the remaining 45 words, only 40 are common to both lists.

Considering the first 500 words on both lists, 86 per cent of the function words, but 58 per cent of the content words are held in common. And even using Otto and Chester's criterion of the percentage of the 500 Atlantic and Pacific words appearing on the first 1000 words of the Newman-Bailey list, the commonality of content words is still only 75 per cent, while that of function words is 91 per cent.

In the first part of the present study, the percentage overlap among a number of other lists was determined. Two hundred words were chosen as a suitable basis for comparison. We found that the first 200 words on a frequency list contain about 35 per cent function and 65 per cent content words. This is a sufficient number from which to assess their relative commonalities. This size list was also chosen because it seems a good size for a sight vocabulary. Finally, the 200 most frequent words in a corpus will account for about two-thirds of the total words (Hillerich, 1974).

The first two lists we compared have already been described, i.e. the Atlantic and Pacific, and Newman-Bailey. Third is the Johnson (1971) list, which is based mainly on the Kucera-Francis (1967) analysis of written materials for adults. Johnson eliminated those words from the Kucera-Francis list which did not occur with a certain frequency in a study of the oral language of Kindergarten and grade one children conducted by Murphy (1957). The result is a 306 word list, of which the first 200 were chosen for the present comparison. Two lists derived directly from the Murphy studies were also included; one is the 200 most frequent words from the Murphy kindergarten corpus, and the other is the 200 most frequent words from the Murphy corpus of grade two children's speech. The two Dolch lists were combined to form another list. The final list is the Hillerich (1974) list which consists of the 190 words common to several lists based on written materials, among them the Kucera-Francis list.

In comparing these lists, all invlected words were collapsed to their lexical forms. Thus, if ask, asks, and asked appeared separately on a list, only the root word ask was included. This was done for three reasons. First, Hillerich argues that this is

the proper form for words on a sight word vocabulary list. If a child can read the word ask, it should not be too difficult to teach him the word asked. Secondly, one use of a sight word vocabulary is the preparation of materials for children to read. It would seem unreasonable to constrain an author with regard to the various forms that words might take. If, for example, an author decides to write a story about baking, because that verb occurs frequently in children's spoken language, it is not practical to disallow the use of the past tense of that verb -- baked -- because it might be a low frequency item. Thirdly, it seems unreasonable to say that two lists differ because they contain different forms of the same word.

However, in preparing a lexical list, irregular forms of a root word are tabulated separately. Thus, although bake and baked are collapsed, run and ran are not, neither are catch and caught. In this study, the following regular endings were removed from words in order to collapse the inflected forms into a lexical list: the regular plurals -s, -es, and -ies; the diminutive -y; third person plural of the verb -s, and -es, regular past tense -ed; present progressive -ing; past participle -en, noun person ending -er, contractions, adverbial form -ly; comparative and superlative adjective forms -er and -est, adverbial forms -y, -ly, and -ily; adjective to noun form -ness, and noun to adjective form -ful. Compound nouns were not broken down into their constituent parts, i.e. snowman was not reduced to snow and man. Appendix A gives examples of the use of each of these endings.

Combining inflected forms reduces the Hillerich list from 240 to 236 words. Although Hillerich's list was also lexical,

he defined root forms in a slightly different way. The Dolch list was reduced to 302 items from 315. The remaining lists were constructed so as to contain exactly 200 lexical items.

From 32 per cent to 40 per cent of each list was composed of function words. Table 1 gives the percentage of function words in common for each pair of lists. The level of agreement is quite high, but varies somewhat by type of list. There are four lists derived mainly from written materials (Atlantic and Pacific, Johnson, Dolch, and Hillerich), and the six comparisons among these lists average 95 per cent agreement. Comparisons among the three oral lists (Murphy Kindergarten, Murphy Grade Two and Newman-Bailey) average 90 per cent agreement, and the twelve comparisons of oral vs. written lists average 81 per cent. Thus, the levels of agreement for function words is quite high, but there do seem to be noticeable differences among lists derived from written and those based on oral materials.

Table 2 gives the levels of agreement for content words. Among written lists, the average is 72 per cent. For oral lists, an intermediate value of 57 per cent is obtained, and the average agreement for written vs. oral lists is 47 per cent. Thus, there seem to be important differences among lists in content words, even among those of the same type.

A Study of the Vocabulary of Toronto Children

The above findings suggest that important differences may exist even among lists based on oral language. Therefore, the second part of the study was undertaken to construct vocabulary lists which reflected the word usage of contemporary Toronto children.

Four different groups of Toronto children were selected because it was thought that their needs might differ.

TABLE 1

COMPARISON OF VOCABULARY LISTS ON FUNCTION WORDS
(Per Cent of Words in Common)

Type[a]	Lists	Johnson	Dolch	Hillerich	Murphy Kinder.	Murphy Grade 2	Newman-Bailey (73 words)
WRITTEN							
	Atlantic & Pacific (77 words)	100%[b]	95	97	77	83	88
	Johnson (82 words)	-	90	94	74	80	85
	Dolch (78 words)	-	-	92	76	85	88
	Hillerich (80 words)	-	-	-	75	80	84
ORAL							
	Murphy Kindergarten (63 words)	-	-	-	-	84	94
	Murphy Grade 2 (67 words)	-	-	-	-	-	93

a "Written" lists are those based mainly on an analysis of written materials. "Oral" lists are based on speech.

b The percentage is calculated by placing the number of words in common over the total size of the list indicated by the row.

TABLE 2

COMPARISON OF VOCABULARY LISTS ON CONTENT WORDS
(Per Cent of Words in Common)

Type[a]	List	Johnson	Dolch	Hillerich	Murphy Kinder.	Murphy Grade 2	Newman-Bailey (127 words)
WRITTEN							
	Atlantic & Pacific (123 words)	74[b]	71	84	45	53	58
	Johnson (118 words)	-	68	88	42	48	52
	Dolch (234 words)	-	-	47	37	44	44
	Hillerich (156 words)	-	-	-	38	46	54
ORAL							
	Murphy Kindergarten (137 words)	-	-	-	-	53	58
	Murphy Grade 2 (133 words)	-	-	-	-	-	59

a "Written" lists are those based mainly on an analysis of written materials. "Oral" lists are based on speech.

b The percentage is calculated by placing the number of words in common over the total size of the list indicated by the row.

One representative elementary school was selected from each of the following types:

(1) Canadian - Advantaged: a school with children from English-speaking backgrounds of high socio-economic status.

(2) Canadian - Disadvantaged: a school with children from English-speaking backgrounds of low socio-economic status.

(3) New Canadian (Integrated) - Disadvantaged: a school with children from a variety of non-English-speaking backgrounds and low socio-economic status.

(4) New Canadian (Segregated) - Disadvantaged: a school with children from a non-English-speaking, mainly Portuguese, background and low socio-economic status.

A study of school 1 allows us to see the level of English language skills that primary school children might develop under the most favourable conditions. School 2 will show the effect of socio-economic status on vocabulary and language development. A study of school 3 will show the effect, if any, of coming from a non-English speaking background, and school 4 will show the effect of the particular composition of the school peer group. Children in school 3 might have more exposure to English than children in school 4 since the variety of ethnic backgrounds present in school 3 precludes the use of any other language as a lingua franca among the children. In school 4; however, children from an ethnic background might speak less English since they have another language in common among themselves.

In each school, grade 1, 2 and 3 teachers were asked to collect oral stories or monologues from a sample of their students. An attempt was made to include 50 children at each grade level from each school. A total of 457 children were actually obtained.

Teachers were instructed to be as non-directive as possible in stimulating the children to talk. The purpose was to collect a sample of speech that was as natural as possible. It was for this reason that teachers rather than research assistants were asked to do the data collection. We felt that the children would talk more freely to someone with whom they were familiar.

Four levels of "directiveness" were discussed with the teachers. At the first level, teachers were encouraged to simply ask the child to "tell a story," giving him or her complete freedom in choosing the topic. It was hoped that most children would respond at this level.

For any who did not, the teacher could then move on to the next level and ask a question which was still quite general, but focused the child's thoughts a bit more -- e.g., "Can you tell me about something interesting that happened last week?" If the child still did not talk freely, the teacher could prompt the child by selecting a topic she felt might interest him or her: the family, a pet, a trip. Finally, the teacher might prompt the child by asking about a very specific event, a TV show, or a book the child had read.

Once a child had selected a topic, the teacher was to encourage him or her to talk at length and elaborate. General positive feedback, such as "uh-huh," and "I see," were to be used to stimulate the child to continue on a topic. General prompts, such as "what happened next?" and specific prompts relating directly to the content were also to be used.

Teachers were asked to attempt to obtain 5 minutes of running speech from each child, i.e. with pauses and hesitations not counted. In most cases, several topics had to be introduced in order to produce this amount, and the teacher likely had to use various levels of prompts at different times during the discourse.

There has been criticism that studies of child language, which are generally done in the classroom, force the child to speak a dialect of English which is not native to him in the presence of teachers from a different social background. There is some evidence

that such procedures may inhibit the child's fluency and richness of expression (Labov, 1970: Houston, 1970).

Without denying the validity of these criticisms, we chose to do our study in the same way for two reasons, one theoretical and one practical. The first reason is that it is precisely a child's classroom English which is of concern. It is this register which he or she will be producing regularly for the teacher and on which the teacher must build. This is that idiolect of his language which is closest to standard English, and it is standard English that the child must eventually learn to read if there are going to be any materials upon which to practice his skill. For the children from ethnic backgrounds, the classroom may be the only situation in which they speak any English at all.

Our study then represents a compromise between those who feel that children should first learn to read in their native language or dialect (Wolfram, 1970) and those who believe that only standard English should be found in the classroom. This study is based on the assumption that children should learn to read standard English, but that a teacher should start from that subset of standard English vocabulary and syntax which the child already controls.

A second reason for choosing to sample classroom speech is that this is what other studies have done. Since one of the purposes of this research is to see how different the needs of Ontario children may be from those of other English-speaking children in other times and places, it is necessary to collect comparable samples of their speech.

Another consideration is whether spontaneous or elicited language should be used. Spontaneous speech would produce language which is most "natural." However, it would be heavy with dialogue and restricted in content, stimulated as it would be by the particular events occurring when it was produced. Such speech is well suited to a study of syntax. However, speech elicited by the teacher through a set of non-directive questions can produce language on a wide range of topics, a feature which is essential to the study of vocabulary. It will also produce speech heavy with monologue, a feature which is also a characteristic of written stories.

The tape-recorded stories obtained through these procedures were transcribed directly into a 360 IBM computer using the ATS system. In compiling the transcriptions, no attempt was made to reproduce a phonoligically accurate record. Rather, all utterances were reproduced in conventional English spelling. Thus "gonna" becomes "going to," "ya," becomes "yes," "-in" becomes "-ing," etc. Contractions; however, were transcribed as spoken -- e.g., don't, it's, can't -- since it is acceptable to present them as such in print. Our decisions in spelling were governed by what was permissible in print.

Some utterances were eliminated from the text. Short non-productive answers by the child to teacher questions were not recorded, e.g., "yes," "no," or "I don't know." Filled pauses were also eliminated, e.g, "um," "ah," or "oh." Repetitions and self-corrections were likewise eliminated. Thus "I want a...a boat." would appear as "I want a boat," and "I go...went to Ontario Place." would be recorded as "I went to Ontario Place."

We also noted where in the text the teacher had prompted the child and where the child spontaneously changed topic. Sentences were defined by falling intonation. Unintelligible words were marked by "X" and unintelligible strings of words by "XX."

The total corpus analyzed in this way was composed of 186,151 words, or an average of 407 words per child. These "stories" were produced using an average of 14 teacher questions or prompts per child. The stories averaged 16 unintelligible words or groups of words. The computer system was used to count the number of times that each inflected item occurred. Inflected forms were then collapsed into their root lexical forms, and the combined frequency of with which the root word occurred was calculated. Finally, lists were constructed for each group consisting of the 200 most frequently occurring lexical items.

Analysis of the Toronto Lists and Comparison with Previous Lists

Each of the four Toronto lists so generated contained 35 per cent function and 67 per cent content words, which is similar to the breakdown of the non-Toronto lists analyzed above. There was likewise a very high level of agreement in the function words contained in the four lists. All comparisons showed an overlap of 94 per cent (see Table 3).

The overlap among Toronto lists in content words was 76 per cent across all comparisons (see Table 3). This is slightly higher than the average overlap of 72 per cent obtained among non-Toronto lists based on written materials, and considerably higher

than the 57 per cent overlap found among oral lists or the 47 per cent overlap found between oral and written lists.

It is interesting that there should be such high levels of agreement among the vocabularies of such disparate groups of children. Referring to Table 3, it can be seen that agreement is highest (80 per cent) between the two groups whose mother tongue is English. But the agreement among the other pairs is only a few percentage points less.

TABLE 3

COMPARISONS AMONG TORONTO LISTS
(Per Cent of Words in Common)

List	Canadian Advantaged	Canadian Disadvantaged	New Canadian (Integrated)	New Canadian (Segregated)
	Part A: Function Words			
Canadian-Advantaged (65 words)	-	94	94	94
Canadian-Disadvantaged (64 words)	-	-	94	94
New Canadian (Integrated) (65 words)	-	-	-	94
New Canadian (Segregated) (65 words)	-	-	-	-
	Part B: Content Words			
Canadian-Advantaged (135 words)	-	80	73	74
Canadian-Disadvantaged (136 words)	-	-	75	78
New Canadian (Integrated) (135 words)	-	-	-	76
New Canadian (Segregated) (135 words)	-	-	-	-

Table 4 gives the percentage overlap for function words between each of the non-Toronto and the four Toronto lists. The overall level of agreement is high (93 per cent) for both non-Toronto lists based on written materials and those based on speech. Only the Murphy Kindergarten list shows substantially less agreement with the Toronto lists. The non-Toronto lists show equally high levels of agreement with each of the four Toronto populations.

Table 5 gives the percentage agreement between each of the non-Toronto and the four Toronto lists for content words. The overall average is 57 per cent. Again the Murphy Kindergarten list is lower than the others (44 per cent). The Dolch list is higher, reaching 70 per cent. Thus, although it is the oldest, the Dolch list appears to be as good as any.

In comparing these lists, it is interesting that the Toronto oral lists showed relatively little agreement with either other oral lists or with written lists. This is consistent with the previous comparisons among non-Toronto lists which also found little agreement among oral lists. However, the Toronto lists did agree fairly well among themselves.

These findings cannot be explained by differences of history and location. The two Murphy lists were compiled during the same period of time and in the same place, but they, too, fail to agree very well. The findings cannot be explained by differences in the age of the children, because the Murphy Kindergarten list agrees only slightly better with the Newman-Bailey preschool list than it does with the Murphy Grade 2 list. If differences in time, place and age were operating together, we would expect the Toronto lists to

TABLE 4

COMPARISONS AMONG TORONTO & NON-TORONTO LISTS
ON FUNCTION WORDS
(Per Cent of Words in Common)

List	Canadian		New Canadian	
	Advantaged (65 words)	Disadvantaged (64 words)	Integrated (65 words)	Segregated (65 words)
Atlantic & Pacific (77 words)	98	98	94	95
Johnson (82 words)	98	98	97	98
Dolch (78 words)	97	95	92	95
Hillerich (80 words)	97	97	95	95
Murphy Kindergarten (63 words)	78	81	78	82
Murphy Grade 2 (67 words)	92	94	92	92
Newman-Bailey (73 words)	95	97	94	97

TABLE 5

COMPARISONS AMONG TORONTO & NON-TORONTO LISTS
ON CONTENT WORDS
(Per Cent of Words in Common)

List	Canadian		New Canadian	
	Advantaged (135 words)	Disadvantaged (136 words)	Integrated (135 words)	Segregated (135 words)
Atlantic & Pacific (123 words)	59	55	52	55
Johnson (118 words)	55	51	49	49
Dolch (234 words)	72	73	70	74
Hillerich (156 words)	65	61	59	59
Murphy Kindergarten (137 words)	45	48	44	49
Murphy Grade 2 (133 words)	57	57	56	55
Newman-Bailey (127 words)	58	60	57	57

show lower levels of agreement with the other oral lists than the oral lists do among themselves. But this is not the case.

It seems most reasonable that the pattern of results is due to differences in the procedures used in the various studies. The two Murphy and the Newman-Bailey lists were all based on studies in which pictures, films, and stories were shown to the children in order to stimulate their thinking and aid the free-flow of speech. Since a large set of different stimuli were used, different children talked about different topics and used different vocabulary. The frequency lists derived from these studies may better reflect the frequency with which various topics occurred in the materials shown the children than they reflect the topics which children like to talk about.

In our study; however, we tried to be as non-directive as possible and elicit speech related to the child's own experience. The relatively high levels of agreement, even for content words, suggests that children do not differ very much when talking about their own lives. Reading through the samples, one hears about home and families, pets and games, learning activities in school and field trips. Perhaps these do not vary a great deal in our society, even among groups from supposedly vastly different backgrounds.

It thus seems that the possibility of constructing a sight word vocabulary for widespread use is fairly good, if it is based on the oral language of children talking about themselves. But the published lists currently in existence are not very appropriate.

A Toronto List

Since there is a relatively high level of agreement among the four Toronto lists, it seems reasonable to construct a sight word vocabulary based on them.

This was done by including any word which appeared on two or more of the lists. This produced a Function Word list of 67 items, and a Content Word list of 137 items. The two lists together, which appear in Appendix A, total 204 items.

A Comparison of the Vocabulary in Beginning Readers With the Toronto and non-Toronto Lists

Olson (1965) analyzed the vocabulary of 7 different readers, and found that only 12 per cent of the words were common to 5 or more of the series. Thus, we might expect that some readers at least would not agree very well with the vocabulary lists we have examined. We selected for examination six reading series which are in fairly wide usage. They are as follows:

(1) The Ginn Integrated Language Program, Levels 1 and 11. Author: M. Kambeitz. Publisher: Ginn & Co. Toronto, 1969. This series includes the following books: What a Dog!, Meet my Pals, Take a Peek and Up the Beanstalk. The series contains word lists which are found in the "To the Teacher" sections of each book. Both Core Vocabulary and Enrichment items were included. When inflections are removed, there are 48 function and 133 content words.

(2) The Nelson Language Development Reading program, Primary level (K-3). Author: John MacInnes. Publisher: The Nelson Company: Toronto, 1970. This series includes: Surprises, Kittens and Bears, Whiskers, and Toy-Box. It contains 71 Function and 178 content words. These words are also listed in "To the Teacher" sections.

(3) The Chandler Reading Program. Authors: D.M. Baugh and M. P. Pulsifer. Publisher: Noble and Noble, New York, 1957. Books included are: Words to Read, and six paper-back readers, Swings, Slides, Trucks, and Cars to Ride, Bikes, Supermarket, and Let's Go; also Let's See the Animals, and Let's Take a Trip. Word lists are included in the back of each book. These word lists contain 65 function and 217 content words.

(4) The Scott Foresman Reading System. Authors: I. E. Aaron et al. Publisher: Scott, Foresman and Company: Glenview, Illinois, 1971. Nine level-2 books were included. Words were taken from text, since there were no reading lists. They include 45 function and 177 content words.

(5) ita - Phase 1. Authors: H.J. Tanyzer and A. J. Mazurkiewicz. Published: Initial Teaching Alphabet Publications, Inc.: Toronto, 1966. included: Ready for Reading, Rides, Dinosaur Ben, Houses, Books 2 and 3 and the Work-books. Word lists are given in the Teacher's Manual and Keys, and the Master List for the I/T/A vocabulary cards. There are 51 function and 194 content words.

(6) Open Court Basic Readers, first half of grade one. Author: A. Hughes. Publisher: Open Court Publishing Company. Books included are: Learning to Read and Write and Reading and Writing. Word lists found in the back of the books included 34 function and 333 content words.

The same type of analysis was performed, with the percentage overlap for function and content words determined between each series and each of the Toronto and non-Toronto lists (see Tables 6 & 7).

Agreement among function words averages 88 per cent. Agreement between each of the six series and the non-Toronto lists is slightly higher (91 per cent) than agreement with the Toronto lists (83 per cent). Both ita and the Ginn series agree unusually well with the Toronto lists, while the Open Court series shows a lower level of agreement.

Levels of agreement for content words are similar to previous comparisons (an average of 50 per cent). Some series do better than others. The Chandler series is high, with an average of 65 per cent agreement with Toronto lists. Chandler is followed closely by Nelson with 62 per cent agreement. Ginn averages 54 per cent, and ita is only slightly lower at 50 per cent. Open Court again is relatively low (42 per cent), followed by Scott-Foresman (39 per cent).

Thus, on the basis of vocabulary, the series can be evaluated for use with Toronto populations as follows: High - Chandler and Nelson; Medium - Ginn and ita; low - Open Court and Scott-Foresman. This ordering is based primarily on the findings for content words. For function words, the high and medium categories would be reversed. The low category; however remains the same.

A Comparison of the Toronto Samples on Grammatical Measures

In the final part of the present study, a limited investigation was conducted of the linguistic complexity evidenced by children from the four Toronto groups. The first two measures are the average story length per group and the type-token ratios. The latter is the number

TABLE 6

PER CENT AGREEMENT BETWEEN READING SERIES AND VOCABULARY LISTS ON FUNCTION WORDS

Lists	Ginn (48 words)	Nelson (71 words)	ITA (51)	Open Court (34 words)	Scott-Foresman (45 words)	Chandler (65 words)
Atlantic & Pacific (77 words)	96%[a]	89	98	89	87	95
Johnson (82 words)	98	92	100	100	89	97
Dolch (78 words)	98	93	98	91	91	97
Hillerich (80 words)	98	92	100	100	91	97
Murphy Kindergarten (63 words)	88	79	90	74	80	83
Murphy Grade 2 (67 words)	92	87	90	80	87	82
Newman-Bailey (73 words)	98	85	98	80	91	94
Canadian - Advantaged (65 words)	88	80	90	74	82	82
Canadian - Disacvantaged (64 words)	92	80	92	74	80	86
New Canadian Integrated (65 words)	92	80	90	71	82	82
New Canadian Segregated (65 words)	92	79	88	77	82	83

a Percentages were calculated using the number of function words in the column list as denominator since, in general, the column lists were smaller.

of different words in a text divided by the total number of words. It is a measure of vocabulary richness.

As Table 8 shows, there was no systematic variation among the four groups on either of these measures. The most productive children (i.e. "talkative") were those in the Canadian-Disadvantaged group. They also had a high type-token ratio. The Canadian-Advantaged and New-Canadian Integrated groups were slightly less productive and had less rich vocabularies. The New-Canadian Segregated group; however, was considerably less productive, but had a relatively rich vocabulary.

We also adapted a method developed by Morehead and Ingram (1973) for analyzing clinical samples from individuals. We chose sentences from subsamples of each of four new groups of children. From the Canadian-Advantaged group, we eliminated the few children who were not native speakers of English. This formed a purified Canadian-Advantaged group. From the two New Canadian schools, a group of English speakers was selected to constitute a new Canadian-Disadvantaged group. Portuguese children from the two New Canadian schools constituted two new groups: a Portuguese-Integrated and a Portuguese-Segregated group. Sentences were collected at random from approximately the middle of each child's "story," until about 85 sentences had been obtained for each group.

Averages were found for the number of words per sentence and the number of morphemes per sentence. A "morpheme" is the smallest unit of meaning in speech. Some words have only a single morpheme, but many are composed of more than one morpheme. The rules for determining the number of morphemes in a word were those given by Gleason (1967). A few are given below for purposes of illustration.

TABLE 7

PERCENT AGREEMENT BETWEEN READING SERIES AND VOCABULARY LISTS ON CONTENT WORDS

Lists	Ginn (244 words)	Nelson (235)	ITA (194)	Open Court (331 words)	Scott - Foresman (177 words)	Chandler (217 words)
Atlantic & Pacific (123 words)	47[a]	67	44	50	40	70
Johnson (118 words)	45	60	42	47	37	61
Dolch (234 words)	44	61	45	40	35	56
Hillerich (156 words)	42	57	43	45	35	58
Murphy Kindergarten (137 words)	45	58	37	33	32	50
Murphy Grade 2 (133 words)	53	68	46	38	39	65
Newman-Bailey (127 words)	51	70	47	46	45	66
Canadian - Advantaged (135 words)	56	61	50	44	38	65
Canadian - Disadvantaged (136 words)	56	63	50	44	39	66
New Canadian Integrated (135 words)	52	61	51	40	36	64
New Canadian Segregated (135 words)	53	64	50	41	41	65

a Percentages were calculated using the number of content words in the column list as denominator since, in general, the column lists were smaller.

TABLE 8

DEVELOPMENTAL COMPARISONS OF THE 4 TORONTO SAMPLES

Group	Story Length in Words	Type/Token Ratio
Canadian-Advantaged	473	.09
Canadian-Disadvantaged	524	.11
New Canadian Integrated	474	.09
New Canadian Segregated	317	.11

(1) Roots and affixes. The word "walk" has only one morpheme, but "walks," "walked," and "walking" each have two.

(2) Irregular verbs. "Find" has one morpheme because it is simple present tense, but "found" has two, because it has the meaning "past" in addition to the root meaning of "find."

(3) Comparative and superlative adjectives. "Big" has one morpheme, but "bigger" and "biggest" have two.

(4) Plural of irregular nouns. "Man" has one morpheme, but "men" has two. However both the singular and plural of "sheep" have only one morpheme.

The number of words and morphemes per sentence is an index of grammatical complexity. The number of morphemes per sentence has been used by Roger Brown as the primary index for language development (Brown, 1973). Tyack and Gotteleben (n.d.) propose a third measure, the Word-Morpheme Index, which is simply an average of the two averages.

This data, which is given in Table 9, gives an interesting picture of the process of language acquisition among children in Toronto. First of all, we notice a large drop-off of about 43 per cent on all three measures between the Canadian-Advantaged and the Canadian-Disadvantaged groups. This is congruent with the large body of literature showing that lower class children have less adequately developed language ability: e.g., for a review see Raph (1965).

There is an additional small, but consistent drop-off (7 per cent) on all three measures between Canadian-Disadvantaged children and their Portuguese counterparts in the integrated school. There is a further drop (6 per cent) between the Portuguese children in the integrated environment and those in the segregated.

There is thus a cost to being a non-native speaker of English, but it is very slight. There is an additional small cost to living in an ethnic ghetto where there is less pressure to speak English.

Two additional measures show a similar pattern. One of these measures is the number of advanced constructions per sentence. We defined these following Tyack and Gotteleben; examples are given below:

> We are going to guess which one you're going to pick out. (Embedded sentence)
>
> I tell them that I don't like it. (Compliment).
>
> I tried it but the first time I fell down. (Coordinated sentence).
>
> When my father's home, we stay home with him. (Adverbial clause).
>
> And after that they would clean out the barn. (Use of modal auxiliary).
>
> So the next day the big bad wolf... (Use of advanced conjunction).

TABLE 9

DEVELOPMENTAL COMPARISON OF ENGLISH AND PORTUGUESE SAMPLES

Measure	English-Speaking		Portuguese-Speaking	
	Advantaged	Disadvantaged	Integrated	Segregated
Words per sentence	18.3	10.5	9.8	9.2
Morphemes per sentence	21.5	12.1	11.3	10.7
Word-Morpheme index	19.9	11.3	10.5	9.9
Advanced constructions per sentence	1.22	0.60	0.38	0.28
Errors per sentence	0.29	0.36	0.54	0.74
(Number of Sentences)	83	88	93	94

The average number of advanced constructions per sentence shows a substantial drop of 51 per cent between Canadian-Advantaged and Canadian-Disadvantaged children, with a subsequent drop of 37 per cent to Portuguese-Integrated and a further drop of 26 per cent to Portuguese-Segregated.

The other measure is the number of grammatical errors per sentence. The largest category of errors was incorrect tense formation in the verb, as in the following examples:

Yesterday I play with my baby.

He catched the girl, turned her (into a) mouse and after he take it home.

Incorrect verb agreement also occurred with some frequency:

Then my uncle give me five dollars.

There's flags on the beach.

My brother watch TV.

Another large category of errors related to incorrect use of the pronoun, as in the following examples:

My brother he says. (Unnecessary pronoun).

Sometimes ( ) have milk and toast. (Omission of pronoun).

Her teach me how to do it. (Incorrect pronoun).

Other types of errors also occurred:

I found a big sock and it was this lady sock. (Incorrect possessive form).

They trying to steal my hockey cards. (Absence of auxiliary verb).

Those little bottle of things. (Incorrect formation of plural noun).

And she says for me to get out. (Incorrect preposition).

Another picture I see the man. (Omission of preposition).

In the last row of Table 9 we see that the error rate increased from an average of .29 per sentence for Canadian-Advantaged to .36 for Canadian-Disadvantaged, to .54 for Portuguese-Integrated, and .74 for Portuguese-Segregated.

It is noteworthy that the disadvantaged and New Canadian groups scored as well as the Canadian-Advantaged group in verbal productivity (story length) and vocabulary richness (type-token ratio). These groups are as able to use language to communicate a variety of ideas as are advantaged Canadian children. However, they have less command of the subtleties of English grammar.

In the case of disadvantaged children who are native speakers of English, we can expect, based on other research, that this initial lack of language ability will continue throughout their school life, resulting in a lower level of overall educational achievement.

However, with New Canadians, it is known that their eventual educational attainment is not inferior to that of native Canadians, nor is there as strong a relationship between socio-economic status and school success (Wright, 1970). Thus, although the two Portuguese groups appear to be economically disadvantaged, and are new speakers of the language, there is a good chance that they will overcome their problems with English and do well in school. In time, they will learn the grammatical rules of English that they now do not know.

New Canadian children in schools where there are many speakers of their native language may need special encouragement to speak English. However, the overwhelming import of the data is that low social class is a much more restrictive factor in attaining a high level of English language ability than is language background, even given a high level of ethnic segregation.

Summary and Abstract

This study has investigated the agreement between previously published vocabulary lists recommended for use with children and lists derived from the speech of contemporary grade 1, 2, and 3 children in Toronto.

high levels of agreement were found among previously published lists for function words - i.e. prepositions, conjunctions, pronouns, forms of the "be" verb, and verb auxiliaries. This is not surprising since the number of function words in English is limited. However, there was less agreement for content words - i.e. nouns, verbs, adjectives, and adverbs. Agreement among lists based on written materials was nevertheless still quite high. Agreement among

lists derived from different studies of children's speech was considerably less. Such lists did not agree very well with lists based on written materials, nor with other speech-based lists, including those derived from the speech of Toronto children.

The four Toronto populations that were studied differed widely in socio-economic and ethnic background. However, agreement among these lists for content words was considerably higher than agreement of the Toronto lists with other oral lists or agreement of other oral lists among themselves. This is probably due to a difference in the procedures followed in this and previous studies.

These results suggest that: 1) the oral vocabulary of children is quite different from material written for children and 2) the oral vocabulary of children is quite sensitive to the manner in which it is collected. However, given procedures like ours that encourage children to talk about their own experience, it seems that different groups of children will use quite similar vocabulary.

This latter finding is relevant to other studies which have been done on the language of disadvantaged children. It has been claimed that the testing procedures used in most studies elicit an artificial and non-representative sample of the children's speech. If tested in a truly natural situation, the argument goes, the speech of disadvantaged groups would be equivalent to that of more economically advantaged groups.

In the present study, it was found that economically advantaged and disadvantaged children, as well as those from non-English backgrounds, were equally talkative and had equally rich vocabularies. Thus, we have confidence in the validity of our procedures. However, on some other measures of language development, specifically grammatical

complexity, use of advanced constructions, and avoidance of grammatical errors, group differences did emerge.

It was found that the scores of economically disadvantaged children were much lower than those of advantaged children. The scores of disadvantaged children from non-English speaking backgrounds were slightly lower still, followed by the scores of non-English speaking children from an ethnically segregated school. This was probably because the latter group had more opportunity to use their native language in school and had less practice in using English. However, the differences due to socio-economic status far outweighed the effects of language background.

The vocabulary of the four Toronto groups was also compared with the vocabulary in six beginning reading series. For function words, all series showed very high levels of agreement with children's speech. However, only moderate levels of agreement were obtained for content words, showing higher levels of agreement than others.

Given the high level of agreement found among the four Toronto lists, it seems that the possibilities of constructing a sight word vocabulary for widespread usage are quite good. A new vocabulary list based on the Toronto study is included in the report, and is suggested as a suitable list for use with a wide variety of children beginning to read.

REFERENCES

Brown, R. A first language: the early stages. Cambridge, Mass.: Harvard University Press, 1973.

Carroll, J. B. A new word frequency book Elementary English, 1972, 49, pp. 1970-1974.

Chester, R. & Otto, W. Children's sight words - printed materials or oral production. The Journal of Educational Research, 1974, 67, pp. 247-252.

Dolch, E. W. Teaching Primary Reading. Champaign Ill.: The Gerrard Press, 1941, pp. 253-258.

Gleason, H. A. Introduction to Descriptive Linguistics (Rev. ed.), New York: Holt, Rinehart & Winston, 1961.

Harris, A. J. & Jacobson, M. D. Some comparisons between basic elementary reading vocabularies and other word lists. Reading Research Quarterly, 1973-74, 9, pp. 87-108.

Hillerich, R. L. Word lists - getting it all together. The Reading Teacher, 1974, 27, 353-360.

Houston, S. A re-examination of some assumptions about the language of the disadvantaged child. Child Development, 1970, 41, pp. 947-963.

Johnson, D. D. A basic vocabulary for beginning reading. The Elementary School Journal, 1971-1972, 72, pp. 29-34.

Kucera, H. & Francis, W. N. Computational analysis of present-day American English. Providence, R.I.: Brown University Press, 1967.

Labov, W. The logic of nonstandard English. In J. E. Alates (Ed.), Monograph Series on Languages and Linguistics: 20th Annual Round Table. Washington, D. C.: Georgetown University Press, 1970.

Morehead, D. M. & Ingram, D. The development of base syntax in normal and linguistically deviant children. Journal of Speech and Hearing Research, 1973, 16, pp. 330-352.

Murphy, H. et al. The spontaneous speaking vocabulary of children in primary grades. Journal of Education, 1957, 140, pp. 1-105.

Newman, M. A. & Bailey, B. E. A compilation of preschool children's word frequency counts. Unpublished manuscript, 1971, (ERIC Document Reproduction No. ED 074 083).

Olson, A. V. An analysis of the vocabulary of seven primary reading series. Elementary English, 1965, 42, pp. 261-264.

Otto, W. & Chester, R. Sight words for beginning readers. Journal of Educational Research, 1972, 65, pp. 435-443.

Raph, J. B. Language development in socially disadvantaged children. Review of Educational Research, 1965, 35, pp. 389-400.

Tyack, D. & Gotteleben, R. Language sampling, analysis and training procedures. Unpublished manuscript, n.d.

Wolfram, W. Sociolinguistic alternatives in teaching reading to non-standard speakers. Reading Research Quarterly, 1970, 6, pp. 9-33.

Wright, E. N. Student's background and its relationship to class and program in school (The Every Student Survey). Toronto: The Board of Education for the City of Toronto, Research Department, 1970 (#91).

APPENDIX A

A Toronto Vocabulary List

The following lists are lexical in nature, that is, the various regularly inflected forms of a word were collapsed to the basic or root form. This includes the following inflections.

- regular noun plurals: -s, -es, -ies, as in "balls," "watches" and "babies"
- diminutives: -y, as in "daddy"
- regular verb third person singular: -s, -es, as in "runs" and "catches"
- regular verb past tense: -ed, as in "looked"
- regular verb present progressive: -ing, as in "walking"
- regular verb past participle: -en, as in "broken"
- regular change of verb to noun: -er, as in "player"
- contractions: -n't, -'re, -'s, -'ll, -'ve, as in "won't," "we're," "it's," "I'll" and "I've"
- comparative and superlative: -er, -est, as in "bigger" and "biggest"
- regular change of adjective to adverb: -y, -ly, -ily, as in "runny," "beautifully" and "prettily"
- regular change of adjective to noun: -ness, as in "fullness"
- regular change of noun to adjective: -full, as in "beautiful"

For some items on the list, selected inflected forms were given in parentheses. These are meant to be suggestive of the productivity of the list, rather than exhaustive. We have not included grammatically incorrect forms which the children may have used - e.g., "eated," or "breaked," - although we counted such forms in deriving the list.

Function Words

a
about
after
and
another
are
around
at
be
because
before
but
by
can (can't)
could (couldn't)
did (didn't)
do (don't, doing)
does (doesn't)
down
for
from
going
had (hadn't)
has (hasn't)
have (haven't, having)
he (he's, he'll)
her (hers)
him
his
I (I'll, I'm, I'd, I've)
if
in
into
is (isn't)
it (it's, its)
let (let's)
me
my
of
off
on
or
our (ours)
out
over
she (she's, she'll)
some
something
that (that's)
the
their (theirs)
them
these
they
this
to
up
us
was (wasn't)
we (we're, we'll, & we've)
were (weren't)
what (what's)
who
with
would (wouldn't)
you (you're, you'll)
your (yours)

Content Words

again
all
always
away
baby (babies)
back
bad (badly)
ball
bed
big (bigger, biggest)
boat
book
boy
break (breaks, breaking)
broke
brother
call (calls, called)
came
car (cars)
cat
Christmas
clean (cleans, cleaned, cleaning, cleaner)
come (comes, coming)
cousin
dad (daddy)
day
dog
door
eat (eats, eating, eaten)
every
father
fell
fight (fights, fighters)
first
fish (fishes, fishing, fished)
four
friend (friends, friendly)
fun
game
gave
get
girl
give
go (goes)
good (goodness)
got
guy
help
here (here's)
hit
hockey
home
house
how
jump (jumps, jumped jumping, jumper)
just
know (knows, known)
lady (ladies)
last (lasted)
like
little
live
look
lot (lots)
made
make
man
mom (mommy, momma)
more

morning
mother
much
name
never
new
nice
night
no
not
now
old (older, oldest, olden)
once
one
only
other
outside
park (parks, parked, parking)
people
place
play
put
ride
right
room
run
said
saw
say
school
see
sister
sleep
so
sometime
start
stay
story
stuff (stuffed)
swim (swims, swimming)
take
tell
then
there
things
think
three
time (times, timer)
told
too
took
toys
try (tries, tried)
T.V.
two
use
want
watch
water
way
well
went
when
where
won
work
year
yes